A Novel Optical Design for Light Field Acquisition Using Camera Array

Mei Zhang *[a], Zheng Geng [a], Zhaoxing Zhang [a],Xuan Cao [a]
[a]The State Key Lab. of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China

ABSTRACT

There is pressing need for 3D imaging technology in many areas. A number of lightfield camera designs are proposed using single image sensor. However, due to the limited size of image sensor chip and optical design, the disparity of the lightfield captured using single sensor camera systems is very small. Stanford group pioneered an implementation of lightfield capture systems using camera array. But, since the camera array often employs discrete imaging sensors and associated optics, the coverage image area for 3D reconstruction is limited. We propose a novel optical design approach that customizes the design for each optical channel to maximize the image quality, coverage area, among other design targets. We then integrate the optical design of all imaging channels into a single monolithic piece with compact structure, high reliability and assembly precision. As a result, the captured light field images from all imaging channels have the same object size with uniform image quality, thus greatly improve the quality of 3D light field reconstruction.

Keyword: optical design, 3D imaging acquisition, light-field capture, camera array, aspherical surface, 3D technology, image system, 3D image reconstruction

1. INTRODUCTION

We propose a novel optical design approach that incorporates an integrated optical design to optimize all optical parameters for all sensors towards a common goal. Instead of design and use the same optics for each and every sensor in the array, we customize the design for each optical channel to maximize the image quality, coverage area, among other design targets. We also integrate the optical design of all imaging channels into a single monolithic piece, incorporating it with structural components to minimize the overall footprint of the camera array, and enhance system reliability and assembly precision. The captured light field images from all imaging channels have the same object size with uniform image quality, thus greatly improve the quality of 3D light field reconstruction.

*mei.zhang@ia.ac.cn; phone 86 13511056416; fax 86 10 82544554

Light field capture techniques have received considerable attention in recent research. In 1990s, Les [1] and Lippmann [2] first discovered the approaches to obtain light field information by an array pinholes and micro lenses, respectively. Subsequently, Adelson [3] described the light field with seven-dimensional parameterization. Later, Levoy [4] simplified the plenoptic function expression as four-function light field parameterization. A number of light field camera designs based on different optical devices are studied. Comparing traditional cameras, a light field camera placed special optical elements in different locations to complete 4D light-field acquisition [5]. Several light-field cameras were proposed using single image sensor. Lee et al. [6] put a biprism in front of the lens for stereo vision with a single camera. Georgeiv et al. [7] applied an array of lens-prism pairs in front of the main lens to capture light field. It has equal function with an array of lenses. Lytro and RayTrix [8-11] implemented lightfield cameras based on placing lenticular optics in front of image sensor. But, the spatial resolution for gaining angular is significantly loss. MIT and MERL [12,13] groups place optical components on aperture plane to manipulate system's point spread function and light modulation mechanism . This technique enables to capture multiple images with full resolution using single photo to estimate depth. However, due to the limited size of image sensor chip and optical design, the disparity of the lightfield captured using these single sensor camera systems is very small.

Stanford group [14] pioneered an implementation of lightfield capture systems using camera array. Wilburn et.al. [15] developed miniature optics to finish light field capture. Some other researches devoted themselves to improve the performance of the lightfield capture system. Cao et al. [16] worked on to reduce camera numbers and propose a dictionary –based light field acquisition. Zhang et al. [17,18] developed an integrated real-time 3D image acquisition and multiview 3D display system. The physical displacement among multiple cameras can be large thus eliminating the problem of small disparity in lightfield capture. However, these light field capture systems based on camera array architecture often employ discrete imaging sensors and associated optics. The optical performance and structure integrity are compromised.

2. OVERVIEW OF OUR APPROACH

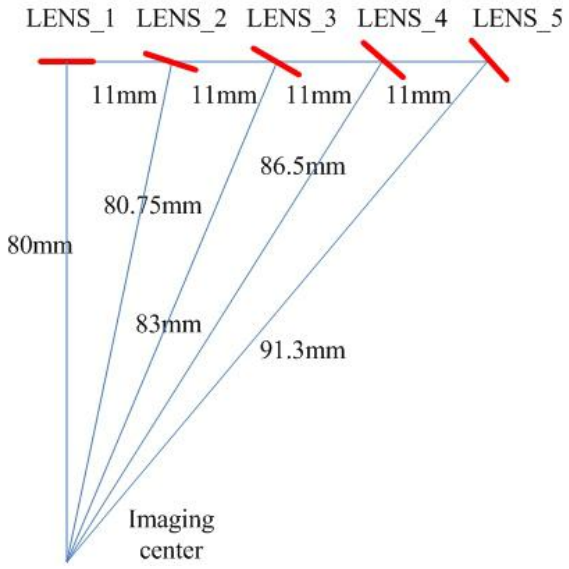


Figure. 1 The alignment of camera array

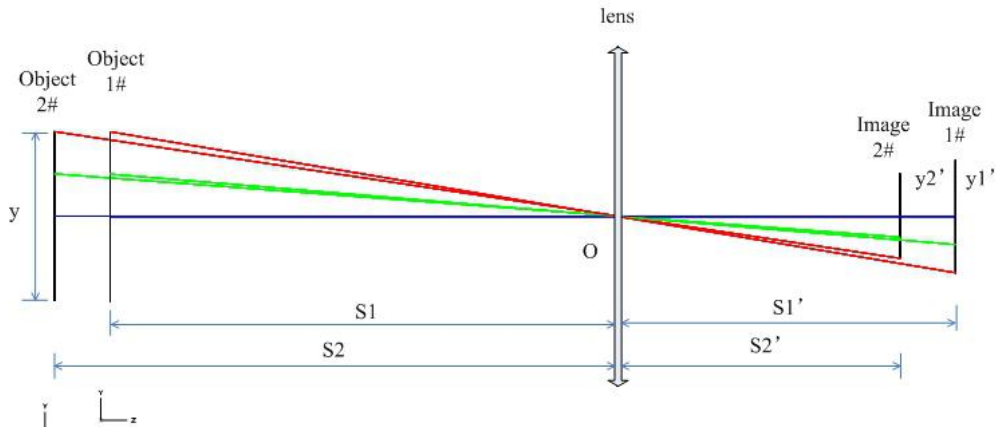


Figure. 2 The different magnification caused by the variation of objective distance for a camera

Generally, several identical cameras are always adjusted on a line or a planar surface to obtain light field information for infinite scene. For the case of a capture system with relatively short standoff distance, all cameras are always directed to a common conformal imaging center to obtain a large overlapping FOV. We take an example of five cameras with equal distance of 11mm, while the conformal imaging center is set 80mm away, as shown in Figure1. The light-field disparity between each camera is about 7.8° . We can calculate that he camera at the edge (LENS_5) is about 91.3mm away from the imaging center, while the first camera (LENS_1) is only 80mm away from the center. According to the optical fundamental, the magnification and the objective distance are directly related. As we can see from Figure 2, when the object is moved from location 1# to 2#, the image surface goes closer and become smaller. The mathematical framework is generally expressed as followed:

$$\frac{1}{s}+\frac{1}{s'}=\frac{1}{f} \tag{1}$$

$$\frac{y'}{y}=\frac{s'}{s} \tag{2}$$

Where s and s' represent the objective distance and image distance, respectively. The y and y' represent the object height and image height, respectively. Let f donate the focal length of the camera.

As shown in Figure2, we take s_1 to be M times of f, and s_2 is equal to N times of f. M and N are given to be much larger than 1 to maintain the image distance nearly unchanged, which means that s_1' is nearly equal to s_2'. From the equation (1) and (2), we can calculate: $\frac{y_1'}{y_2'}=\frac{s_2}{s_1}=\frac{N}{M}$. Comparing LENS_5 with LENS_1, $\frac{y_1'}{y_2'}$ is equal to80/91.3. Therefore, fewer pixels of overlapping area on CCD are covered. We can estimate that only 80% overlap area for LENS_5 can be used to complete 3D surface reconstruction. And the resolution of the reconstruction surface is reduced.

Our goal is to make an integrated optical design for all cameras to hold common parameters, such as a uniform magnification. It is important to increase the image resolution and coverage area for 3D reconstruction. Although some camera has zoom capability, it is hard to realize for very small size restriction.

We propose an optical design method to built light field camera array with nine miniature cameras on a line. Half part of the symmetrical arrangement with 5 cameras is shown in Figure 1. The objective distance for every image capture channel is different which related with their position. The object surface given as 160mm×160mm is imaged to fulfill the CCD sized 1.2mm ×1.2mm for all cameras. Therefore, all cameras must to be modified to have a consistent magnification value.

3. OPTICAL DESIGN

3.1 Initial Objective Optical System

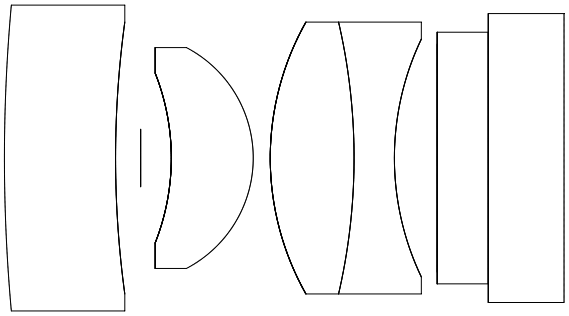


Figure 3. The configurations of initial objective optical system

We designed an optical system for the LENS_5 on the edge to meet the requirement of object height with 160mm. So the field of view is near 81°, which is large enough for our system. Applying the 1.7mm diagonal size of CCD, the effective focal length can be estimated by the equation shown as follow:

$$h' = f \times \tan \omega \tag{3}$$

Where ω is half of the field-of-view and h' is half of the diagonal size of CCD. We can calculate that the focal length f is about 1 mm. To ensure the image quality for such large field-of-view, an inverse telephoto optical system is always used, as shown in Figure 3. Four lenses and two protective planar glasses are involved. The frontal group with negative focal power consists of a concave lens, while the rare group with positive focal power consists of a convex lens and a cemented lens. When off-axis light transmits through the frontal group, the angle formed by the light and axis will reduced greatly, so the angle of the field of view in the image space is much smaller than in the objective space. This character is beneficial to the aberration correction so as to acquire large aperture as well as large field of view. The main parameters of the objective optical system are listed in Table 1.

Table 1. Numeric value data of the initial objective optical system

No	Radius (mm)	Thickness (mm)	Nd	Vd
OBJ	INF	91.3	-	-
1	10.0	0.653	1.52	64.12
2	5.93	0.147	-	
Stop	INF	0.180	-	-
4	-1.364	0.481	1.75	52.33
5	-0.735	0.1	-	-
6	1.632	0.492	1.88	40.81
7	-3.561	0.237	1.92	18.89
8	1.628	0.251	-	-
Filter	INF	0.3	1.52	64.12
Cover	INF	0.445	1.52	62
IMG	INF			

We simulate and analysis the optical system in zemax soft ware. When the objective distance is set as 91.3 mm, the image semi-diameter gives the value of 0.844mm. But when the objective distance is set as 80mm, the image semi-diameter increased to 0.917 mm with about 18% failed to be captured on CCD. Therefore, only about 82% common area of LENS_5 can be matched with LENS_1.

Figure 4 is the MTF of the system. The spatial cutoff frequency dependents on the pixel size of CCD sensor, which is 2μm×2μm in our system. The cutoff frequency is expressed as:

$$f_{cut} = \frac{1000}{2 \times p} \tag{4}$$

Put p=2μm into the equation (4), we can know the spatial cutoff frequency f_{cut}=250lp/mm. From Figure4, the MTF below 125 lp/mm is above 0.35, while it is above 0.15 at 250 lp/mm

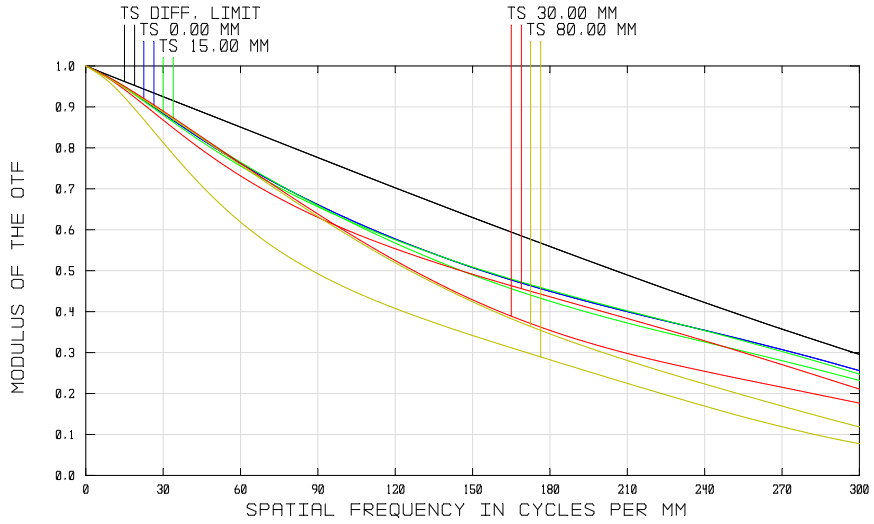


Figure 4. The MTF of the objective optical system

3.2 Integrated Camera Array Design

4.2.1 Integrated Design Method

The main goal of the design is to make the nine optical channels to hold a uniform magnification. After completing initial objective optical design, we need to make an integrated optical design for the whole camera array. The image quality for every optical channel needs to be equally good. Considering the low cost and easy fabrication to realize integration for different inclination angle between cameras, plastics material can be used for lens.

The optical design procedure is described as follow:

(1) Designing an imaging optical system meet the requirement of the system in ZEMAX, as shown in Figure 3.

(2) Setting five configurations for different objective distance requirement, following the data shown in Figure 1.

(3) Setting both surface of the first lens different for every configuration while making other structure parameters multi-used.

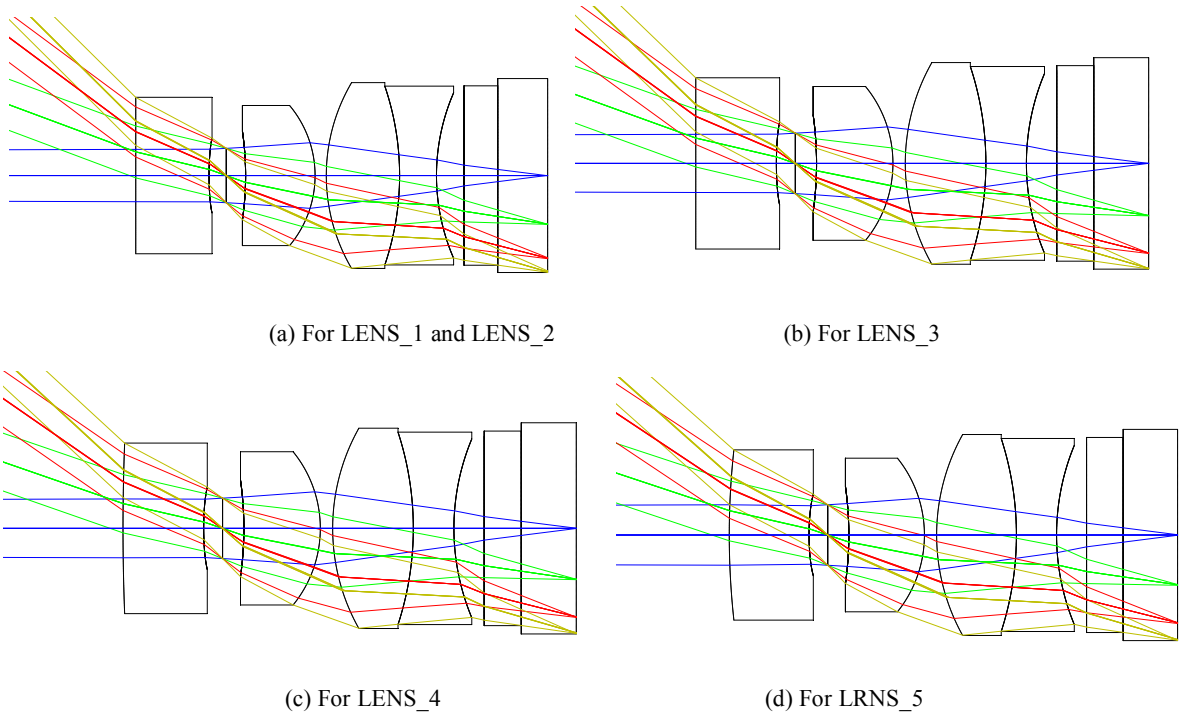
(4) Setting the second surface as aspheric surface to get better performance, and changing the first lens as plastics material to make easy manufacture for such miniature optical lens. The aspheric surface can be mathematically expressed as:

$$z = \frac{cr^2}{1+\sqrt{1-(1+k)c^2r^2}} + \beta_1\gamma^1 + \beta_2\gamma^2 + \beta_3\gamma^3 + \beta_4\gamma^4 + \bullet\bullet\bullet \tag{5}$$

Where c, k, and γ represent radius, coefficient of taper and normalized radial coordinate, respectively, while β_1, β_2, β_3 and so on indicate coefficient of high order aspheric surface.

(5) Making optimization with unified image size for the whole optical multi-configuration system.

Because the difference of objective for LENS_1 and LENS_2 is very small, no need to modify the cameras for them. So we make integrated optical design for LENS_1, LENS_3, LENS_4 and LENS_5. The final configurations of the four cameras are shown in Figure 5. Obviously, only the first lens has different curvatures, while other lenses share same parameters. Table 1 and Table 2 give the parameters in detail. The native power of the first lens becomes stronger and stronger, which reflects that the effective focal length becomes longer and longer. It can be seen that the semi-diameter of image surface for all cameras are very close to each other, around 0.85mm to fully match the CCD size. Considering the difficulty of process, only the 4^{th}, 6^{th} and 8^{th} high order coefficients of the aspheric surface are used. The total length of the integrated camera becomes a little longer (about 3.66mm) than that of the initial objective optical system (about 3.29 mm).



(a) For LENS_1 and LENS_2 (b) For LENS_3

(c) For LENS_4 (d) For LRNS_5

Figure 5. Four configurations of the integrated designed camera array

Table 2. Numeric value data of the integrated camera array

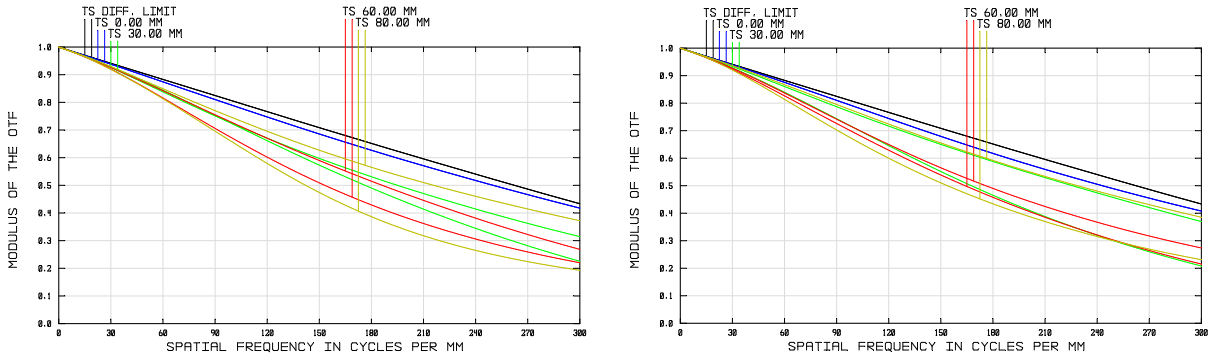
No	Radius (mm)	Thickness (mm)	Nd	Vd
OBJ	INF	OB	-	-
1	R1	0.653	1.49	57.33
2	R2	0.153	-	
Stop	INF	0.172	-	-
4	-2.287	0.619	1.75	52.33
5	-0.975	0.1	-	-

6	1.639	0.653	1.88	40.81
7	-2.483	0.328	1.92	18.89
8	1.862	0.2245	-	-
Filter	INF	0.3	1.52	64.12
Cover	INF	0.445	1.52	62
IMG	INF			

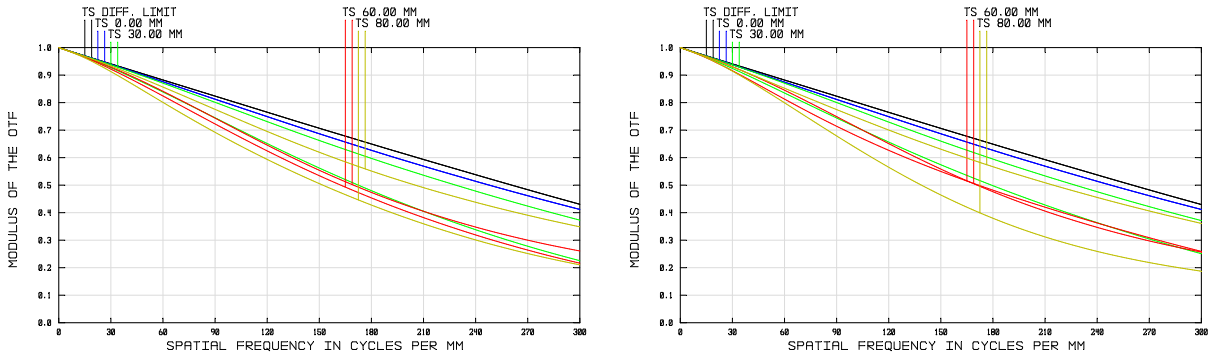
Table 3. Different parameters between cameras

No	LENS_1	LENS_3	LENS_4	LENS_5
OB	80	83	86.5	91.3
R1	INF	INF	17.694	6.190
R2	3.007	2.947	2.541	1.975
Conic	60.097	54.726	39.424	6.821
4th Order	0.34	0.171	0.071	0.194
6th Order	-3.192	0.447	3.005	3.972
8th Order	13.896	-8.000	-23.672	-8.050
EFFL	1.2311	1.2340	1.2479	1.2782
Semi-IMG	0.86	0.85	0.845	0.846

4.2.2 Image performance



(a) (b)

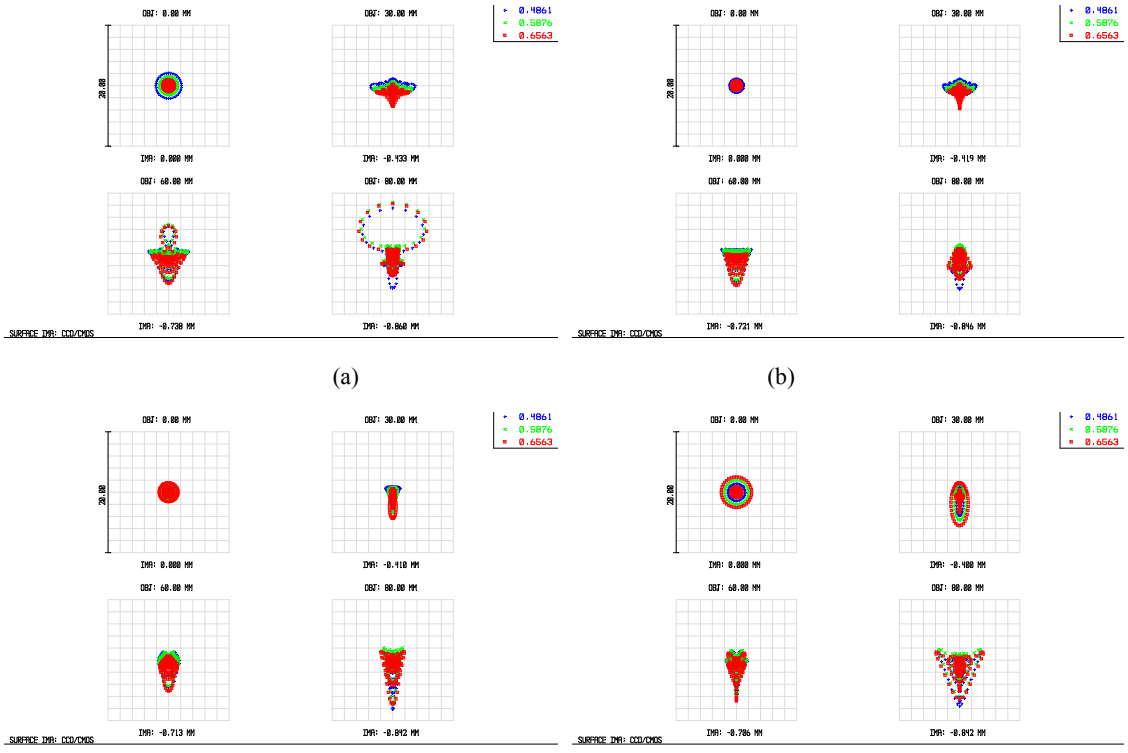


(c) (d)

Figure 6. The MTF curves of the four configurations

To evaluate the image quality comprehensively, MTF curves of every configurations are given in Figure6. The MTF curves seem very likely, and are all above 0.3 at 250 lp/mm for every configuration. This means all aberrations are well corrected. Comparing with Figure 4, the value is increased. That's because aspheric surface is introduced.

The spot diagram for the four configurations are given in Figure 7. The RMS radius for center field-of-view is around 1μm, while for magical field is not more than 2.72μm. In a short, the integrated optical design realizes uniform magnification and good image performance for all cameras without introducing too much fussy modifies.



(a) (b)

(c) (d)

Figure 7. The spot diagram of the four configurations for camera array

Due to the plastic metirial of the first lens, it is easy to fabrate two ears for the first lens. It is helpful to set up nine camaras accoding to their different tilte angle. All nine cameras can be fix into a tube with compact structure, high reliability and assembly precision, as shown in Figure 8.

Figure 8. The mechanical model of the integrated camera array

4. IMPLEMENTATION

In order to reveal the performance of the integrated camera array, we take images to make 3D reconstruction using the initial camera array and integrated designed camera array, respectively. The Structure from Motion (SFM) and Bundle Adjustment (BA) algorithms are implemented in the calibration process to minimize the calibration errors and ensure quality of reconstructed 3D model.

We consider the algorithm DTAM (Dense Tracking and Mapping) [19]to process our images. DTAM is a system for real-time camera tracking and reconstruction which relies not on feature extraction but dense, every pixel methods. The algorithm use whole image alignment against a given dense model to track camera motion at frame-rate, and expand the model by refining dense textured depth maps. Generally, no feature-based skeleton or tracking is required.

A set of 3D dense surface reconstruction algorithms are developed to recover high resolution 3D surface profile of the scene, while preserving feature details and discontinuities. An optimization energy function is designed to convert the inherent non-convex problem into two separate optimization processes that can be solved in alternative steps using duality theory.

(a) (b)

Figure 9. Image variation between central and side cameras with no modify

Figure 9 gives the images taken by the central and side cameras of the initial camera array. Obviously, the objective size

captured by side camera is larger than that by central camera. The feet of the dolls appear in Figure 9 (a) but not in Figure 9 (b). The common area used for 3D reconstruction has to be decreased to fit all nine images. As shown in figure 10, the image captured by side camera is cut down to about 80%, only about 520×520 pixels are covered for common area (600×600 pixels for full CCD). Figure 11 shows the 3D reconstruction results. The feet of the dolls cannot show.

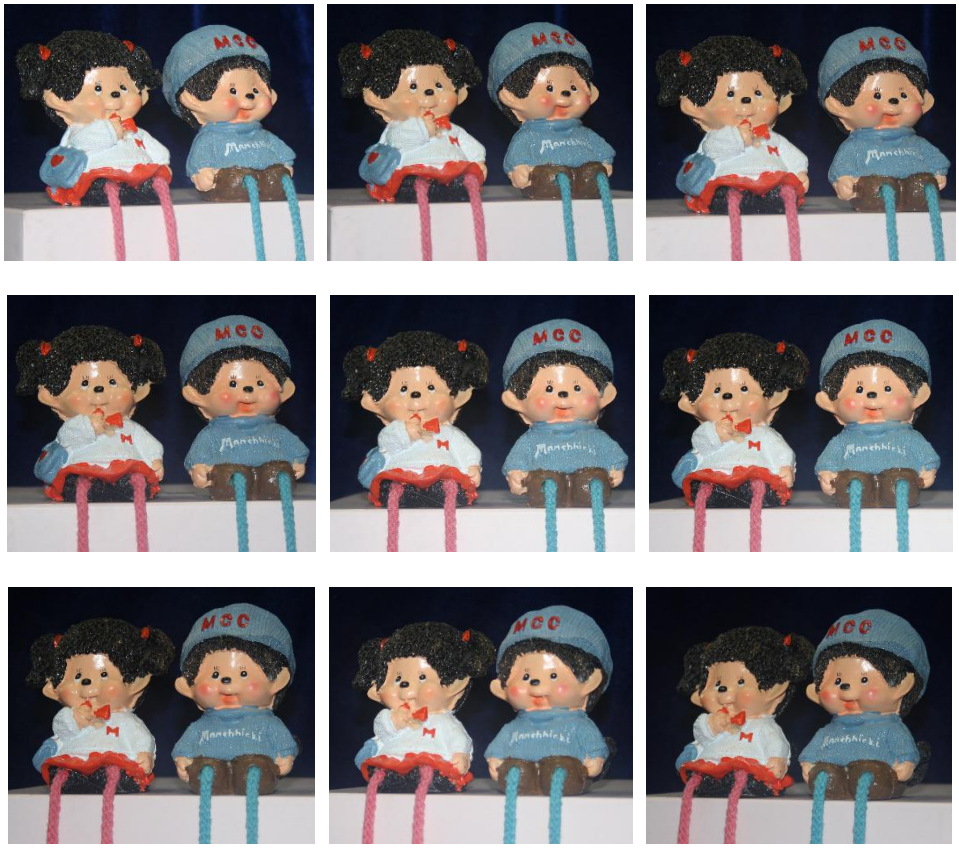


Figure 10. Common area of the nine images captured by initial cameras array

Figure 11. 3D reconstruction results of the nine images captured by initial cameras array

Figure 12 gives the nine images captured by the integrated designed cameras array. The image size nearly remains the

same among the nine images. No more waste area we need to cut. Figure 13 gives the 3D reconstruction results. Comparing with Figure 11, the dolls with feet are shown. More object information can be used to finish 3D reconstruction.

Figure 12. Nine images captured by integrated designed cameras array

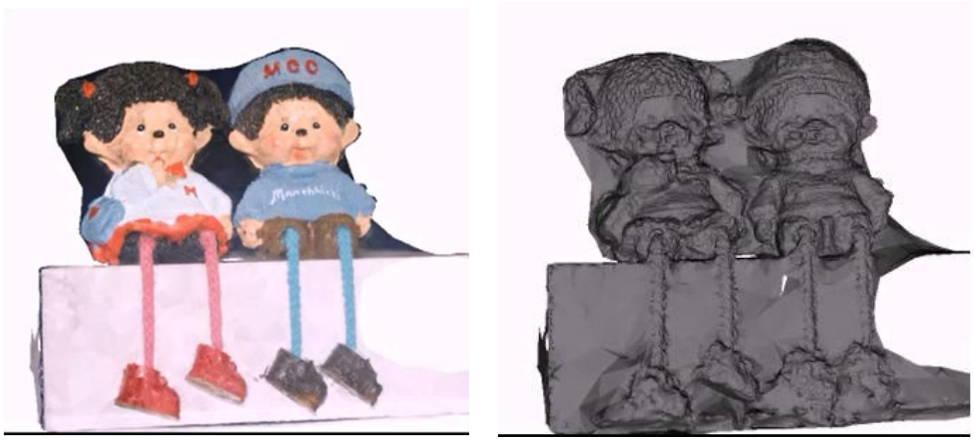


Figure 13. 3D reconstruction results of the nine images captured by integrated designed cameras array

We compare the effective area of 3D reconstruction of the two types of camera array mentioned above. Table 4 gives the result. For common camera array system, the image area of every optical channel extracted for 3D reconstruction becomes less and less. 100% image area of the center camera LENS_1 can be used for reconstruction, while only 82% coverage area of LENS_5 can be processed. However, every optical channel for integrated designed camera array can provide effective image area for reconstruction by one hundred percent. Obviously, the integrated camera array system greatly avoids the loss of image information and improves the performance of 3D reconstruction.

Table 4. Comparison of the effective area of 3D reconstruction for common camera array and integrated designed camera array

	Common camera array	**Integrated camera array**
LENS_1	100%	100%
LENS_2	99%	100%
LENS_3	94%	100%
LENS_4	88%	100%
LENS_5	82%	100%

5. CONCLUSION

Most camera array system aligns discrete identical cameras to make 3D image acquisition, which introduce considerable diversity in image quality and structure integrity problems. We propose an integrated design method to realize a camera array with uniform high performance. Firstly, we successfully designed an integrated camera array with nine cameras. Only the first lens has different geometrical parameters between each other, while other structure parameters are same. Furthermore, we analyzed the performance of the camera array. All cameras with different location can capture the image with identical magnification and excellent image quality. The MTF values for all cameras are all above 0.3 at 250 lp/mm. Finally， we take images and process 3D reconstruction by use of the initial common camera array and integrated designed camera array, respectively. After comparing the results for the two kinds of camera arrays, we can see that the FOV for the integrated camera array is larger than that for common camera array. This work can significantly improve the performance 3D imaging acquisition with camera array to fulfill 3D surface reconstruction with high quality.

Acknowledgments

This work has been supported by the National High-tech R&D Program (863 Program) of CASIA (Chinese Academy of Sciences, Institute of Automation), Grant No. 2012AA011903.

REFERENCES

[1] Levs, F., "US patent 725,567, " (1903).
[2] Lippmann, G.,, "Épreuves réversibles donnant la sensation du relief," J.Phys. Théor. Appl. 7, 812-825 (1908).
[3] Adelson, E.H. and Bergen, J.R., "The plenoptic function and the elements of early vision, " in Computational

Models of Visual Processing, M.S. Landy and J. A. Movshon, eds (MIT Press), pp. 3-20(1991).

[4] Levoy, M. and Hanrahan, P., "Lightfield rendering, " in proceedings of ACM SIGGRAPH, pp.31-42 (1996).

[5] Zhou, C., "Computational cameras: convergence of optics and processing," IEEE transaction on image processing Vol.20,no.12, pp.3322-3340 (2011)

[6] Lee, D., Kweon, I. and Cipolla, R., "A biprism-stereo camera system," in Proc. IEEE Conf. Comput. Vis. Patten Recognit.Vol.1,pp.82-87(1999).

[7] Georgeiv, T., Zheng, K.C., Curless, B., Salesin, D., Nayar, S., and Intwala, C., "Spatio-angular resolution tradeoff integral photography", in Proc. Eurograph. Symp. Rendering, pp.1-10 (2006).

[8] Chen, S. ,Donoho, D. L., Saunders, M. A., "Basis Pursuit," in 28th Asilomar Conf. Signals, Systems Computers, (1995).

[9] Tibshirani, R., "Regression shrinkage and selection via the lasso," J. Royal. Statist. Soc B vol. 58, no. 1, pp. 267-288 (1996).

[10] Elad, M., [Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing], Springer, (2010).

[11] Mallat, S. G. and Zhang, Z., "Matching Pursuits with Time-Frequency Dictionaries," IEEE Transactions on Signal Processing vol. 41, no. 12, pp. 3397-3415 (1993).

[12] Aharon, M., [Overcomplete Dictionaries for Sparse Representation of Signals], Ph.D. thesis, Technion - Israel Institute of Technology, (2006).

[13] Marwah, K., Wetzstein, G., Bando, Y., and Raskar, R., "Compressive Light Field Photography using Overcomplete Dictionaries and Optimized Projections," in SIGGRAPH, (2013).

[14] Efron, B.; Hastie, T., Johnstone, I., Tibshirani, R., "Least Angle Regression," Annals of Statistics vol. 32, no. 2, pp. 407-499(2004).

[15] Wilburn, B., Joshi, N., Vaish, V., Talvala, E., Antunez, E., Barth, A., Adams, A., Horowitz, M., and Levoy, M., "High performance imaging using large camera arrays," ACM Transactions on Graphics Vol.24, no.3, pp.765-776, (2005).

[16] Cao, X., Geng, Z., and Li, T. , "Dictionary-based light field acquisition using sparse camera array," Optics Express Vol. 22, Issue 20, pp. 24081-24095 (2014)

[17] Zhang, Z., Geng, J., Li, T., Li, W.; Wang, J. and Liu, Y., "Integration of real-time 3D image acquisition and multiview 3D display," Proc. SPIE Vol. 8979, (2014).

[18] Geng, J., "Three-dimensional display technologies," Advances in Optics and Photonics Vol. 5, Issue 4, pp. 456-535(2013).

[19] Newcombe, R.A., Lovergrove, S.J. and Davison, A.J., "DTAM: dense tracking and mapping in real-time," IEEE International Conference on Computer Vision, ICCV, Nov.6-13, pp.2320-2327(2011).